

Reply comments in FCC Public Notice on closed-captioning rules

These reply comments pertain to the FCC's 2010 Public Notice (CG 05-231; ET 99-254) on closed-captioning rules. The focus here – again – is captioning *standards*.

Permanent location

This submission, which follows up on my original comments, is permanently located at joelclark.org/access/crtc/fcc2010/reply/. (You're better off reading the online version instead of the printout submitted to the FCC's filing system.)

Of course we'll differentiate between live and prerecorded captioning

I see two themes in the this proceeding's responses to the proposition that the FCC adopt captioning quality standards.

1. **High-level managers seem to think captioning *is* real-time captioning.** I assume this is due to the fact that few, if any, such respondents really watch TV, which they're too busy for and which is kind of tacky and lowbrow anyway. And they certainly don't watch TV with captioning on all the time. Instead, the use case seems to be this one: CNN is on all day on a flatscreen out by the reception desk, with captioning turned on because sound is turned off. Managers walk by this monitor a few times a day, and that's all the captioning they see (not really "watch").

I admit the foregoing is a supposition, but it's been my experience that network executives neither watch a lot of TV nor watch all their TV with captioning. I rather dispute how they can provide expert opinions on captioning under any circumstances. (Why do we listen to people who *do not watch captioning*?)

2. Real-time-captioning houses and operators are **unnerved by the prospect of having *another* thing to worry about** – "standards" – while they're developing carpal-tunnel syndrome captioning college football matches and podunk local newscasts all day.

I understand their alarm, but they need to relax. Real-time and prerecorded captioning are cousins, not identical twins, and would always be treated separately in any proceeding. Real-time captioning can be credibly approached in a statistical manner, for example, while offline captioning cannot.

- On this subject, Caption Colorado's homespun and ad hoc definitions of "readability" are all well and good but are not really supported by existing research into psychology of reading, little of which pertains to the unique task of reading

captions or the unique subtask of reading real-time captions. In particular, Caption Colorado's self-proclaimed Formula for Calculating Readability Rating needs to be taken for what it is – a sincere suggestion from nonexperts that has not been subjected to testing. In fact, the whole “formula” is undermined by the averages Caption Colorado wants captioners to meet. There, the per-transcript averages max out at 97% accuracy while all transcripts as a whole must achieve a mathematically impossible *higher* average, 97.5%. (Is half a percentage point worth talking about in an average?)

- I note that Inclusive Technologies at least is aware that reading research is a specialized field and any such research must be carried out by qualified scientists.

Respondents are disingenuous about broadcasters' true priority in captioning

Corporate respondents swear up and down they have an abiding, day-in/day-out commitment to captioning, but this is disingenuous at best. Broadcasters and program producers shop for captioning on price *and that's it*. Nothing else comes into the picture.

As such, it is an insult to the intelligence to read the National Association of Broadcasters baldfaced claim that broadcasters “strive to ensure that the programming they deliver to their audiences is as error-free as practicable.” What they strive to ensure is that their captioning is as cheap as possible.

NCTA's claims that broadcasters are “continually reviewing the performance of captioning services they use” and that captioners “compete on accuracy” are both false. In the former case, broadcasters don't watch a lot of TV and don't watch TV with captions on all the time. In the latter case, captioners all *claim to be more accurate than everybody else*, with no greater verifiable basis than one brand of lemonade claiming to taste more lemony than all the others.

Media Captioning Services submitted evidence that large blocks of real-time captioning are tendered at zero cost to highly profitable networks like CNN. It has already been attested that NBC hired a contractor based on a reverse or Dutch auction in which the *lowest* bidder won. (That “winner,” CaptionMax, then proceeded to dumb down and eviscerate its captioning to the point where it resembles half-assed subtitling.) Cost is broadcasters' only criterion in selecting a caption house.

Broadcasters want expensive human beings eliminated from the process

For decades, broadcasters' most cherished dream has been to eliminate altogether the costliest part of captioning: People.

Broadcasters did not get into this business to help cripples, and many of them resent having to pay for captioning or just hate captioning outright. Broadcasters are not word people; broadcasting executives tend to be older. Put those two facts together and what you're left with is a picture of a group that really cannot actually watch TV with captioning. They just can't stand it.

It galls broadcasters that they cannot go right ahead and air whatever programming is cheapest, sell commercials against it, and call it a day. Having to arrange for, then pay for, the extra step of captioning sticks in their craw.

The dream they hold dear is as follows: They pull an old computer out of storage, hook it up with some kind of cheap but miraculous transcription software they bought for a one-off price, boot everything up, and leave. Maybe they'd double-check that the computer was still running before they went home for Thanksgiving and Christmas breaks, but that's about it.

The point here is it offends broadcasters that HAL 9000 can't just sit there typing out what people say all day at zero incremental cost. ("I mean, seriously: We're still *paying people* to do this?")

Thus, broadcasters thought their ship had finally come in when various charlatans and snake-oil salesmen showed up at their doors selling "speech recognition" for captioning. No matter how often vendors might qualify the term, broadcasters heard what they *wanted* to hear and thought the holy grail had finally been delivered unto them. Voice recognition... at last, a computer writes it all down for us.

At the consumer level, there is no such thing as a computer that can listen to people talk and transcribe their words. No such computer is going to be invented in our lifetimes. And even if it *were* invented, I remind you that broadcasters' mental image holds that all captioning is real-time captioning. The computer won't just do all the transcribing in this model; it will extrude scrollup captioning in perpetuity. But scrollup captioning is a deficiency when used with fictional narrative programming, and the computer still wouldn't know how to divide, time, and place pop-on captions. It's a non-starter.

What is the reality? Numerous vendors will sell you a system in which *one person* sits in a room and repeats the dialogue from a TV show. The system is trained to understand only that person. From a labour perspective, there is no difference between this voicewriting and real-time stenography; you still need a dedicated person doing it. (But that person sure is cheaper!) As such, broadcasters' dream remains unrealized.

I don't know any knowledgeable parties, other than Martin Block and vendors of these actual systems, who believe respeakers are even remotely as good at captioning as an experienced stenographer. Maybe for slow-moving

events with barely any speech, like golf matches, the viewing experience might be comparable, but for news programming or anything fast and satirical (e.g., *The Daily Show*), forget it.

“Voice recognition” for captioning is really “speaker-dependent voice recognition” for captioning and basically doesn’t work for captioning. Respondents in this process agree, and not just marginal individuals or lobby groups – Caption Colorado, NAB (see below), and especially NCTA (“voice-recognition technology still cannot be relied on as a substitute for live captioning”) concede that voicewriting doesn’t work for captioning.

In a related trend, broadcasters are always telling us the solution to intractable caption problems lies in new technology that’s comin’ right down the pike (conveniently *after* a regulator’s deadline). Well, that technology is *not* coming. You need human beings running computer software to do captioning and that’s that.

To recap, then: Broadcasters hate paying for captioning and really hate paying for the human element in captioning. They want a computer they buy once, then forget about, to do all the work. Some broadcasters think they already have that or have something at least halfway as good as expensive stenography. They don’t.

The National Association of Broadcasters isn’t making sense

Let’s accept that the NAB will always oppose government regulation. It will especially oppose requirements that cost more than nothing. We just sort of accept that, in the way that we accept that Mormon missionaries who knock on our door are going to try to convert us. In neither case do they necessarily get their way, of course.

- NAB contradicts itself. Its submission starts out with *the* clearest statement of the facts about voice-recognition captioning: “[S]peech recognition currently does not match the accuracy level of a real-time stenocaptioner.” Yet it also claims that some kind of “automated captioning” (a leftover computer sitting in a closet somewhere?) might soon be invented, which FCC’s Big Government quality standards would ban:

Broadcasters, MVPDs and others will be deterred from utilizing new technologies that are still improving because of the real possibility of FCC enforcement actions. The inevitable result would be a slowdown in developing and employing these new automated captioning technologies.

Moreover, companies that develop other technologies which, in the future, could be applied to captioning would be discouraged from entering the captioning market because these technologies might not initially meet rigid accuracy benchmarks. Thus, adopting specific accuracy requirements could have the unintended consequence of deterring captioning innovation.

Shorter NAB: If somebody invents an even cheaper new system for automated captioning, we should be allowed to use it even if the captioning sucks.

- NAB's objections to the very premise of a caption-quality standard descend almost to self-parody. If you somehow manage to make it through a full double-spaced page of "reasons" why caption quality is too philosophically intractable to be regulated, all you're left with is a claim that unquantifiable processes *can't* be regulated. Yes, they can.

The fact that standards haven't been written yet ("there is no current agreed-upon industry method for assessing caption accuracy") is no proof that they can't be. That fact is of course *convenient for the NAB*, which wants the status quo maintained so nobody has to spend an extra penny making sure captioning doesn't suck.

- NAB again seems to be fixated on a mental image of real-time captioning standing in for all captioning.
- NAB's claim that verbatim captions would fly past the viewer too fast to be read has been broadly debunked by the research of Jensema (1996, 1998). Even the Brits, who are so clueless about captioning they can't even call it that, concluded there is no a priori basis to limit caption speed.

Complaints are failed mechanism

NAB notes that "the record in this proceeding, including the recent Report on Captioning Informal Complaints, does not [evince] a widespread failure by broadcasters to deliver high-quality captioning." I don't know how many times I need to remind people that it is functionally impossible to complain about the *quality* of captioning of *any particular show*. It's the same amount of work to complain about a "minor" issue of "style" as it is to complain about completely absent or unwatchably garbled captioning. It's a lot of work no matter what you're complaining about, so people don't even bother until things get to the point of missing or garbled captions.

I repeat: Absence of complaints about caption quality tells us nothing about such quality. What it manifestly does *not* tell us is that nobody actually has complaints to lodge. Complaints are too hard to file and are a case of locking the barn door after the horse has bolted. The badly-captioned program already aired. It's history.

The only way to solve the problem is to impose independently-developed and tested captioning standards on everybody *at the outset*. We've been through this before, and I'm still waiting for somebody to make the case that this won't actually work.

As such, Verizon's insistence that "the FCC should focus its finite resources

on handling and pursuing complaints concerning existing closed-captioning practices” will not improve caption quality. Nor do I agree that more regulations are unnecessary because Verizon thinks we have too many already and they aren’t being followed up enough.

“Industry best practices” are a restatement of the problem

Verizon states:

The Commission should encourage the adoption of industry best practices to resolve common closed-captioning problems

[T]he Commission should encourage the video industry to investigate the causes of the common closed-captioning problems that have been identified by TDI, and to adopt industry “best practices” that address those issues.... This targeted, industry-driven approach is far more likely to improve the overall quality of closed captioning than would a broad new set of regulations that increases costs and administrative burdens without looking at the source of the closed captioning problems that exist today.

“Industry best practices” is a euphemism for “no government regulation,” which in turn means “broadcasters continue to shop on price alone.” Industry best practices – really *worst* practices – are what we already have and resulted in captioning that sucks.

Other points

- Caption Colorado’s “proprietary technology” for moving captions to avoid covering up onscreen graphics sounds a lot like adding a user interface to an off-the-shelf data bridge. Procedures like these should be mandatory and not subject to one supplier’s “proprietary technology.”
- The online participants quoted by Inclusive Technologies who call for delaying picture and sound so that captions are perfectly synchronized to them are out of their minds. It isn’t even remotely technically possible, starting with the fact that real-time captioners are often or usually *watching the same show you are* as they caption it.

Don’t blow it, FCC

The Commission has a once-every-half-decade chance to actually improve caption quality. The only option on the table is independently-developed and tested standards. Don’t blow this opportunity.